

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES**

**On Appeal to the Board of
Appeals and Interferences**

Applicant : Paek et al.

Serial No. : 09/830,899 Group Art Unit: 2171

Filed : August 13, 2001 Examiner: Leroux, Etienne Pierre

Title: DESCRIPTION SCHEMES FOR MPEG-7 IMAGE/VIDEO
CONTENTS DESCRIPTION

I hereby certify that this paper is being electronically filed with the United States
Patent Office.

May 16, 2007
Date of Deposit

Robert L. Maier

Reg. No. 54,291

Name



Signature

May 16, 2007
Date of Signature

REPLY BRIEF ON APPEAL

On December 14, 2006, Appellant filed an Appeal Brief in the above-identified patent application opposing the rejection of claims 1-43 presented in the Office Action issued by the U.S. Patent and Trademark Office (the "PTO") on July 3, 2006. An Examiner Answer was mailed on April 5, 2007. In response thereto and in accordance with 37 C.F.R. § 41.41, Appellant submits this brief in support of the appeal of the rejection of pending claims 1-43. For the reasons set forth below, in addition to the reasons set forth in Appellant's December 14, 2006 Appeal Brief, the rejection of pending claims 1-43 should be reversed.

I. STATUS OF THE CLAIMS

In the July 3, 2006 Office Action, claims 1-43 were rejected under 35 U.S.C. § 102(b) as allegedly anticipated by U.S. Patent No. 6,079,566 to Eleftheriadis et al. (hereinafter “Eleftheriadis”). Appellants have respectfully traversed all rejections of record.

II. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

Claims 1-43 were rejected under 35 U.S.C. § 102(b) as allegedly anticipated by Eleftheriadis. In this Appeal, Appellants have respectfully requested review of all rejections of record.

III. ARGUMENT

As an initial matter, Appellants refer back to and incorporate by reference all arguments made in Appellants’ Brief on Appeal submitted in this Appeal on December 14, 2006 (hereinafter the “Appeal Brief”). Appellants respectfully submit that the Examiner’s Answer of April 5, 2007 does not adequately address the numerous deficiencies in the rejections of record, as discussed in the Appeal Brief and further discussed herein.

As discussed in more detail in Appellant’s Appeal Brief, independent claim 1 is directed to a system for generating a description record from multimedia information, comprising:

at least one multimedia information input interface receiving said multimedia information; [e.g., specification, p. 26 (“*Digital image data 710 is applied to the computer system via link 711.*”); pp. 27-28, Fig. 8]

a computer processor, coupled to said at least one multimedia information input interface, receiving said multimedia information therefrom [e.g., specification, p. 26 (“*Digital image data 710 is*

applied to the computer system via link 711.”); pp. 27-28, Fig. 8], processing said multimedia information by performing object extraction processing [e.g., specification, p. 26; Fig. 7, “object extraction 720”; Fig. 3] to generate multimedia object descriptions [e.g., specification, p. 26, “object set 721,” “object descriptions”; Fig. 5] from said multimedia information, and processing said generated multimedia object descriptions by object hierarchy processing [e.g., specification, p. 27; Fig. 7, “object hierarchy extraction and construction module 730;” p. 28; Fig. 8, module 830] to generate multimedia object hierarchy descriptions [e.g., pp. 18-19, 23, 25; Figs. 3, 4a, 4b, 5, 6a, 6b] indicative of an organization of said object descriptions [e.g., pp. 18-19, 23, 25; Figs. 3, 4a, 4b, 5, 6a, 6b], wherein at least one description record including said multimedia object descriptions and said multimedia object hierarchy descriptions is generated for content embedded within said multimedia information [throughout the specification, it is described in numerous instances that the descriptions are generated for multimedia content, e.g., p. 5, line 12; Fig. 3]; and

a data storage system, operatively coupled to said processor, for storing said at least one description record [e.g., Fig. 7, 740; Fig. 8, 840, and related descriptions in specification].

(Claim 1).

Importantly, the claimed subject matter includes the recitation of “performing object extraction processing to generate multimedia object descriptions from said multimedia information, and processing said generated multimedia object descriptions by object hierarchy processing to generate multimedia object hierarchy descriptions indicative of an organization of said object descriptions, wherein at least one description record including said multimedia object descriptions and said multimedia object hierarchy descriptions is generated for content embedded within said multimedia information.”

(Claim 1) (See, e.g., Specification pp. 11-14, 26-29; Figs. 2, 3, 7, 8; and related citations to the specification and drawings as indicated above). Similar recitations are included in independent method claim 17, including, e.g.:

receiving said multimedia information [e.g., *specification*, p. 26 (“*Digital image data 710 is applied to the computer system via link 711.*”); pp. 27-28, *Fig. 8*];

processing said multimedia information by performing object extraction processing [e.g., *specification*, p. 26; *Fig. 7*, “*object extraction 720*”; *Fig. 3*] to generate multimedia object descriptions from said multimedia information [e.g., *specification*, p. 26, “*object set 721*,” “*object descriptions*”];

processing said generated multimedia object descriptions by object hierarchy processing [e.g., *specification*, p. 26; *Fig. 7*, “*object extraction 720*”; *Fig. 3*] to generate multimedia object hierarchy descriptions [e.g., *specification*, p. 26, “*object set 721*,” “*object descriptions*”; *Fig. 5*] indicative of an organization of said object descriptions, wherein at least one description record including said multimedia object descriptions and said multimedia object hierarchy descriptions is generated [*throughout the specification, it is described in numerous instances that the descriptions are generated for multimedia content, e.g., p. 5, line 12; Fig. 3*]

storing said at least one description record [e.g., *Fig. 7, 740; Fig. 8, 840, and related descriptions in specification*].

(Claim 17).

and in independent computer-readable medium claim 33, which includes, *inter alia*:

one or more multimedia object descriptions, generated by performing object extraction processing [e.g., *specification*, p. 26; *Fig. 7*, “*object extraction 720*”; *Fig. 3*], said object descriptions describing corresponding multimedia objects [e.g., *specification*, p. 26, “*object set 721*,” “*object descriptions*”] [*throughout the specification, it is described in numerous instances that the descriptions are generated for multimedia content, e.g., p. 5, line 12; Fig. 3*];

one or more features characterizing each of said multimedia object descriptions;

one or more multimedia object hierarchy descriptions indicative of an organization of said object descriptions [e.g., *specification*, p. 26, “*object set 721*,” “*object descriptions*”; *Fig. 5*], if any, relating at least a portion of said one or more multimedia objects in accordance with one or more characteristics [e.g., *specification*, p. 26, “*object set 721*,” “*object descriptions*”; *Fig. 5*].

(Claim 33).

Grounds for Rejection

The Examiner's Answer raised several new/different allegedly supporting citations in Eleftheriadis that were not cited in the Office Action. Appellants hereafter address those new citations in turn in connection with the claim language as referenced in the Examiner's Answer:

Independent claim 1 recites, *inter alia*:

“processing said multimedia information by performing object extraction processing to generate multimedia object descriptions from said multimedia information”

and

“wherein at least one description record including said multimedia object descriptions and said multimedia object hierarchy descriptions is generated for content embedded within said multimedia information”

In connection with the above-quoted claim language, the Examiner merely provides the following, without any particular citation to any portion of Eleftheriadis:

“[file contains a header having streaming information, physical object information and logical object information]”

The Examiner apparently refers to the Summary of the Invention, as quoted from Eleftheriadis below (at col. 2, lines 7-12):

The invention accomplishing these and other objectives in one aspect relates to a method of composing and extracting data in a file, and to a medium storing that type of file, with **the file containing a header having streaming information, physical object information and logical object information**, and a sequence of audiovisual segments containing audiovisual objects and a segment object data table pointing to access information, to access the audiovisual objects in each audiovisual segment.

If anything, this language would support Appellants' position that the present application and Eleftheriadis relate to two entirely different aspects of multimedia information processing. As in many other instances in the Office Action and in the

Examiner's Answer, the Examiner erroneously conflates the *streaming* of information with the subject matter of the present application.

As explained in the Background of the Invention, the Abstract, and the Field of the Invention of the present application, “[t]he present invention relates to techniques for describing multimedia information... .” (Specification, page 1, lines 1-2). As explained in the Appeal Brief, the present application relates to techniques for describing the contents of multimedia information (for example, video) so that information can be cataloged and easily searched at a later date (for example, in the way that an Internet user can utilize a search engine such as Google.com to search for text information on the Internet). These aims of the invention, and of MPEG-7 generally, are described throughout the Specification. (See, e.g., Specification, pp. 2-4).

The claimed subject matter of the present application is accordingly unconcerned with “streaming” multimedia information. As explained on numerous occasions previously during the prosecution of this application, this is expected given that the present application pertains to, e.g., MPEG-7 multimedia description techniques while Eleftheriadis pertains to, e.g., MPEG-4 encoding/streaming/presentation techniques. The Examiner's continued equating of the streaming of multimedia information in the prior art with the extraction of descriptive information about that multimedia information (as claimed in the present invention) is unfounded.

Examiner's Response to Argument

Appellants respectfully submit that the Examiner has included other similar logical incongruities in the “Response to Argument” portion of the Examiner's Answer.

- 1) “processing said multimedia information by performing object extraction processing to generate multimedia object descriptions”

On page 10 of the Examiner’s Answer, the Examiner alleges that the “object descriptions” of the claimed subject matter are the same as the cited definition for the word “description.” Appellants strongly disagree. One of ordinary skill in the art would clearly understand based on the specification, drawings and claims of the present application that the “object descriptions” of the claimed subject matter have a meaning that is something more particular than the Examiner’s cited definition for the word “describe” in a vacuum.

Moreover, the Examiner argues as follows:

The above teaching that individual segments containing audio-visual objects can be processed, encoded, stored and retrieved reads on the claim limitation “processing said multimedia information by performing object extraction processing to generate multimedia object descriptions.”

Examiner’s Answer, p. 11.

Once again, the Examiner has actually *distinguished* the claimed subject matter from Eleftheriadis. The argument above explains that Eleftheriadis teaches that “individual segments containing audio-visual objects can be processed, encoded, stored and retrieved.” Appellants agree completely with the Examiner’s assessment of Eleftheriadis. As supported by the above-quoted passage, Eleftheriadis relates to video streaming/presentation to a user, and therefore necessarily involves all of the above-cited actions. However, the claimed subject matter above that the Examiner equates to these above-cited actions actually has nothing to do with these actions at all.

The subject matter of the independent claims relate instead to processing the multimedia information by performing object extraction processing to generate multimedia object descriptions. Again, the purpose of the claimed subject matter is to generate object descriptions by receiving a multimedia stream as an input and extracting descriptive

information about the objects from the multimedia stream – to facilitate the cataloging/searching of that multimedia stream later. This is *entirely* distinct from the MPEG-4 encoding/decoding/presentation considerations of Eleftheriadis. As argued previously, Appellants respectfully submit that this distinction would be absolutely clear to one of ordinary skill in the art.

The Examiner also cites the following passage from Eleftheriadis:

The AV (audio-visual) objects making up a scene are separately encoded and stored in file segments, and composition data for composing scenes out of those constituent objects is separately stored and can be randomly accessed and readily edited as well. Moreover the invention is capable of processing MPEG-1, MPEG-2 audio, video and systems data files, along with coded MPEG-4 data with its extended capabilities.
Eleftheriadis, col. 2, lns. 30-35.

The Examiner thereafter emphasizes that the “constituent objects can be randomly accessed and readily edited (emphasis added) reads on the claim limitation ‘processing said multimedia information by performing object extraction processing to generate multimedia object descriptions.’” (Examiner’s Answer, p. 11). However, this cited portion of Eleftheriadis is also entirely unrelated to the claimed subject matter.

The Examiner quotes the following text of Eleftheriadis in support of his position: “composition data for composing scenes out of those constituent objects is separately stored and can be randomly accessed and readily edited as well.” The editing of composition data (i.e., data within a stream that is used to *compose* scenes that include the video objects – in other words, the data required to present the video information to a user) is entirely irrelevant to the present invention’s object extraction processing to generate descriptions of the multimedia information. If anything, the cited portion of Eleftheriadis once again emphasizes that Eleftheriadis is directed to MPEG-4

encoding/decoding/presentation/etc., while the presently-claimed subject matter is instead directed to providing descriptive information about such multimedia information (in accordance with MPEG-7). Appellants again respectfully submit that one of ordinary skill in the art would understand this clear and important distinction.

Next the Examiner explains:

Furthermore, Eleftheriadis discloses in column 5, lines 10-13 that object IDs are used to uniquely identify the AV (audio-visual) objects encapsulated in AL PDUs 60, including the BIFS (binary format scene description information). The above teaching that object IDs are used to uniquely identify the AV (audio-visual) objects reads on the claim element “processing said multimedia information by performing object extraction processing to generate multimedia object descriptions.”
Examiner’s Answer, p. 11.

Appellants do not understand this comparison. The generation of object descriptions, as recited in the claimed subject matter at issue, is unrelated to the use of unique identifiers in Eleftheriadis. These unique identifiers are used to reference objects in the video stream – they do not generate or provide any type of “description” of the objects. Conversely, the object descriptions of the present invention are descriptions of the contents of the multimedia.

The specification describes several exemplary object description techniques. First, the specification explains that “first subsystem 210 is a region-based indexing and searching system which extracts visual features such as color, texture, motion, shape, and size for automatically segmented regions of a video sequence. The system 210 decomposes video into separate shots by scene change detection, which may be either abrupt or transitional (e.g. dissolve, fade in/out, wipe). … For each shot, the description generated by this system is a set of regions with visual and motion features, and the camera motion.” (Specification, p. 11).

Next, the specification refers to a second subsystem 220 which “provides a set of regions with face labels.” (Specification, p. 12).

As a further example, the fifth subsystem 250 is described as used for the entry of manual text annotations, e.g., to capture textual features presented on screen in conjunction with on-line news sources. (Specification, p. 13).

The object descriptions of the claimed subject matter are clearly something more “descriptive” than the unique identifiers of MPEG-4, which do not *describe* the content of the multimedia information. And, in any event, the cited portion of Eleftheriadis still does not address at least the object extraction processing of the claimed subject matter.

Finally with respect to the above-discussed claim language, the Examiner cites to the “one-byte Profile field 460 (Figure 1) containing profile/level descriptions for each AV Object present in the file.” (Examiner’s Answer, p. 12). However, the Profile field of Eleftheriadis does not provide descriptive information about the contents of the multimedia information as do the object descriptions of the claimed subject matter. This cited portion of Eleftheriadis also again does not describe extraction of objects or generating descriptions of those objects. Rather, the citation relates only to data commonly used to encode/decode MPEG-4 multimedia data streams for transmission and/or presentation.

Accordingly, the Examiner has not shown that Eleftheriadis discloses or suggests the claimed “processing said multimedia information by performing object extraction processing to generate multimedia object descriptions.” For at least this reason, Appellants respectfully submit that the rejections of all pending claims should be reversed.

2) “*processing said generated multimedia object descriptions by object hierarchy processing to generate multimedia object hierarchy descriptions*”

On pages 12-13 of the Examiner’s Answer, the Examiner cites several portions of Eleftheriadis as allegedly disclosing the above-cited claim limitations. Appellants respectfully disagree.

The Examiner makes the following assumptions on page 13, allegedly based on the description in the Specification:

- (1) The tree structure of Fig. 6b is a hierarchical organization of objects. The hierarchical organization of objects is plain and simply the sequence of video events, i.e., event 0 is the entire video sequence followed by events 1, 2, 3, 4 and then events 5 and 6.
- (2) There is no difference between, a scene, an event and an image object.

Examiner’s Answer, p. 13

The Examiner thereafter cites the following portion of Eleftheriadis as allegedly disclosing the above-quoted claim language:

Individual components of a scene are coded as independent objects (e.g. arbitrarily shaped visual objects, or separately coded sounds). The audiovisual objects are transmitted to a receiving terminal along with scene description information, which defines how the objects should be positioned in space and time, in order to construct the scene to be presented to a user. The scene description follows a tree structured approach, similar to the Virtual Reality Modeling Language (VRML) known in the art. The encoding of such scene description information is more fully defined in Part 1 of the official ISO MPEG-4 specification (MPEG-4 Systems), known in the art. BIFS information is transmitted in its own elementary stream, with its own time and clock stamp information to ensure proper coordination of events at the receiving terminal.

Eleftheriadis, col. 3, lns. 30-45.

Appellants respectfully disagree with the Examiner’s assumptions, and further disagree that the above-quoted portion of Eleftheriadis discloses or suggests the claimed subject matter.

As discussed in the Appeal Brief, the above-quoted lone reference in Eleftheriadis to a “tree-structured approach” is unrelated to the claimed “processing said generated multimedia object descriptions by object hierarchy processing to generate multimedia object hierarchy descriptions.” As explained in detail in the quoted portion of Eleftheriadis above, that reference refers to methods described in MPEG-4 and, e.g., VRML, for providing multimedia scene information using a tree-structure (for, e.g., streaming video playback/scene presentation). However, the tree structure of Eleftheriadis is entirely unrelated to describing the content of the multimedia information for later search/retrieval. Accordingly, the cited portion of Eleftheriadis bears no relation to the claimed “processing said generated multimedia object descriptions by object hierarchy processing to generate multimedia object hierarchy descriptions.”

The claimed hierarchy, as further described in an exemplary embodiment, e.g., at pp. 17-20 of the present application, relates to an object hierarchy for description of particular video objects with varying levels of specificity – for purposes of content description, and not hierarchy of a scene for composing or presenting the scene or playing back streaming multimedia information. The claimed object hierarchy processing can produce a “physical hierarchy” and a “logical hierarchy,” which relate to the physical location of objects in an image, and a higher level hierarchy based on semantic descriptions of the objects in the image, respectively. (See Specification, p. 17; Fig. 4). The object hierarchy descriptions may include semantic information which is useful for searching a library of multimedia segments, such as “names of the picture, the names of persons in the picture, the location where the picture was taken, the event that is represented by the picture, the date of the picture, color features... .” (Specification, p. 20).

Perhaps most importantly, the hierarchy of Eleftheriadis is not generated by processing “multimedia object descriptions,” as claimed herein. The hierarchy of Eleftheriadis is unrelated to object descriptions. Additionally, the hierarchy of Eleftheriadis does not include “multimedia object hierarchy descriptions.” Once again, these differences are important in view of the aims of Eleftheriadis/MPEG-4 systems and the claimed subject matter/MPEG-7 systems.

Accordingly, because Eleftheriadis fails to disclose or suggest at least these additional claimed features, Appellants respectfully submit that the rejections of record should be reversed as to all claims.

3) *Examiner’s Statement Regarding Inconsistencies in the Specification*

Finally, on page 14 of the Examiner’s Answer, the Examiner refers to “above inconsistencies in the specification.” Appellants are unclear as to what inconsistencies the Examiner refers, or regarding whether those alleged inconsistencies are in the present Specification or the specification of Eleftheriadis. In any event, Appellants are unable to respond to the Examiner’s assertion in this respect due to the lack of any clear description of the alleged issue.

These same arguments apply to all the claim groups. For at least these additional reasons, Appellants respectfully request that the Board reverse the Examiner’s rejection of claims 1-43 under 35 U.S.C. § 102(b) as being disclosed or suggested by Eleftheriadis.

IV. CONCLUSION

For at least the reasons indicated above, in addition to all the reasons set forth in Appellants' Appeal Brief, Appellants respectfully submit that the invention recited in the claims of the present application, as discussed above, is new, non-obvious and useful. Reversal of the Examiner's rejections of the claims is therefore respectfully requested.

Respectfully submitted,

Dated: May 16, 2007

By:



Paul A. Ragusa
Patent Office Reg. No. 38,587

Robert L. Maier
Patent Office Reg. No. 54,291

Attorneys for Appellants
Baker Botts L.L.P.
30 Rockefeller Plaza
New York, NY 10112-4498
Telephone: (212) 408-2500